One Ecosystem 3: e28143
doi: 10.3897/oneeco.3.e28143

OPEN ACCESS

Data Paper

Distribution of bumblebees across Europe

Chiara Polce[‡], Joachim Maes[‡], Xavier Rotllan-Puig[‡], Denis Michez[§], Leopoldo Castro[|], Bjorn Cederberg[¶], Libor Dvorak[#], Úna Fitzpatrick[¤], Frederic Francis[«], Johann Neumayer[»], Aulo Manino[^], Juho Paukkunen[˅], Tadeusz Pawlikowski[¦], Stuart Paul Masson Roberts[ʔ], Jakub Straka[ʕ], Pierre Rasmont[§]

‡ European Commission Joint Research Centre, Ispra, Italy
§ University of Mons, Mons, Belgium
| Unaffiliated, Teruel, Spain
¶ Swedish University of Agricultural Sciences, Uppsala, Sweden
# Municipal Museum Mariánské Lázně, Mariánské Lázně, Czech Republic
¤ National Biodiversity Data Centre, Waterford, United Kingdom
« University of Liège, Gembloux, Belgium
» Unaffiliated, Elixhausen, Austria
^ Università di Torino, Torino, Italy
˅ Finnish Museum of Natural History, University of Helsinki, Finland
¦ Nicolaus Copernicus University, Lwowska, Poland
ʔ University of Reading, Reading, United Kingdom
ʕ Charles University, Prague, Czech Republic

Corresponding author: Chiara Polce (chiara.polce@gmail.com)

Academic editor: Benjamin Burkhard

Received: 03 Jul 2018 | Accepted: 06 Sep 2018 | Published: 13 Sep 2018

Citation: Polce C, Maes J, Rotllan-Puig X, Michez D, Castro L, Cederberg B, Dvorak L, Fitzpatrick Ú, Francis F, Neumayer J, Manino A, Paukkunen J, Pawlikowski T, Roberts S, Straka J, Rasmont P (2018) Distribution of bumblebees across Europe. One Ecosystem 3: e28143. https://doi.org/10.3897/oneeco.3.e28143

Abstract

Insect pollinators are a key component of biodiversity; they also play a major role in the reproduction of many species of wild plants and crops.

It is widely acknowledged that insect pollinators are threatened by many environmental pressures, mostly of anthropogenic nature. Their decline is a global phenomenon. A better understanding of their distribution can help their monitoring and ultimately facilitate conservation actions.

Since we only have partial knowledge of where pollinator species occur, the possibility to predict suitable environmental conditions from scattered species records can facilitate not

© Polce C et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

only species monitoring, but also the identification of areas potentially vulnerable to pollinators decline.

This data paper contains the predicted distribution of 47 species of bumblebees across the 28 Member States of the European Union (EU-28). Amongst the wild pollinators, bumblebees are one of the major groups contributing to the production of many crop species, hence their decline in Europe, North America and Asia can potentially threaten food security.

Predictions were derived from distribution models, using species records with a spatial resolution of 10 km accessed from a central repository. Predictions were based on records from 1991 to 2012 and on a series of spatial environmental predictors from three main thematic areas: land use and land cover, climate and topography.

These distributions were used to estimate the value of pollination as an ecosystem service. In light of the recent European Pollinators Initiative, this paper provides valuable information for a better understanding of where wild pollinators occur and it should be extended to other pollinator species.

Keywords

Bombus, insect pollinators, species distribution model, Maxent, Europe

Overview and background

Pollination is a key ecosystem service vital to the maintenance of both wild plant communities and agricultural productivity. Over three quarters of the world's major crops benefit from insect pollination (Klein et al. 2007), with an annual economic value estimated at $235-577 billion globally (in 2015 US$) (IPBES 2016). In Europe, insect pollination benefits more than 80% of crops (Williams 1994), with an estimated annual value of several billion euros (depending on the methods and crop considered, estimates range from €3 billion to just above €14 billion, see for instance: Vallecillo et al. 2018, Leonhardt et al. 2013, Gallai et al. 2009). Pollination services are mainly provided by wild insect pollinators (bees, hoverflies, flies, moths, beetles etc.) and domesticated bees (primarily western honey bee *Apis mellifera* and *Bombus terrestris*). Hence, there is growing concern that observed declines in insect pollinators, mainly bees (Potts et al. 2010, Biesmeijer 2006, Nieto et al. 2014), may impact on production and revenues from pollinator-dependent crops. Knowing the distribution of pollinators, therefore, is crucial for estimating their availability to pollinate crops (and wild plants). This information, in turn, can be used to ensure the maintenance of habitats that support insect pollinators, ultimately safeguarding the long term provision of (crop-)pollination services. In general, however, there is incomplete knowledge of where wild pollinators occur. To overcome this issue, two main approaches are usually adopted: one based on expert knowledge of the species' ecology (e.g. Zulian et al. 2013), the other based on species records (e.g. Polce et al. 2014, Polce

et al. 2013). Both deliver a relative indicator between 0 and 1 expressing the environmental suitability to support the species of interest. This indicator is often interpreted as probability of occurrence of the species and, hence, used as a proxy for their potential presence.

Each of these approaches has some strengths and weaknesses: an expert-based model (EBM), for instance, can account for the effect of detailed local information, such as the presence of wild flower edges between crop-fields or other small patches of habitat suitable for pollinators. The EBM, however, could fail to reflect the environmental suitability for poorly known species or to capture environmental characteristics that can modify the expected suitability (e.g. climatic differences) or, again, it may not be able to predict species richness. A species-distribution model (SDM), on the other hand, can be formed by actual species records, which are used to characterise the 'quality' of the environment where species are recorded through statistics or machine-learning techniques for instance. An SDM, therefore, is constrained by the spatial and temporal resolution of these records; hence, it could fail to capture the effect of local landscape elements, if their spatial accuracy is greater than that which is available for the species records; or it could lead to biased predictions if the input data are biased (e.g. spatially, temporally biased) and no corrections are applied.

These two approaches have been recently applied to derive a spatial indicator for the 'pollination potential by wild insect pollinators' across the European Union, within KIP INCA project ('Knowledge Innovation Project Integrated system for Natural Capital and ecosystem services Accounting'). The original models by Zulian et al. (2013) and Polce et al. (2013) were adapted to meet the requirements of the accounting, such as the need to rely on datasets that are regularly updated. The EBM was adopted to predict the potential distribution of solitary bees across Europe, while the SDM was used to predict the potential distribution of (wild) bumblebees. The outputs of these models were then integrated to estimate the potential availability of wild insect pollinators across Europe and, finally, the pollination service (Vallecillo et al. 2018). While the EBM largely followed the scores and main dataset from Zulian et al. (2013), the SDM was based on a dataset never used for this purpose, while the approach largely followed Polce et al. (2013). Hence, in this paper, we present the predictions of bumblebee species from the SDM and describe the methods to generate them.

Bumblebees are important pollinators not only of wild plants, but also of crops. So their decline in Europe, North America and Asia is a cause of concern. Like many other species, bumblebees are also sensitive to environmental change. Maps derived from their records across Europe were recently produced to characterise their current climatic niche and their projected distribution based on climate change scenarios (Rasmont et al. 2015) as well as on the interplay between climate and land use change (Marshall et al. 2017). Work of this type allows the understanding of the expected shift in suitable habitat and climatic conditions for each species and, hence, an estimate of the potential risks due to environmental change. Bumblebees are also a highly charismatic group, as they are large brightly coloured insects associated with flowers. Therefore, they can effectively contribute to raising public awareness on the issue of pollinators and engage citizens in pollinators-related activities (e.g. Bumblebee Conservation Trust).

In light of the recent EU Pollinators Initiative, the possibility to predict suitable environmental conditions based on species records is particularly important, not only to facilitate species monitoring over time and across geographical space, but also to predict areas potentially vulnerable to pollinators decline.

Hence, we present the potential distribution of 47 species of European bumblebees, derived from their records and key environmental drivers.

Methods

Overview of the study

For the importance of bumblebees within agricultural production, maps displaying their likely distribution are a key component of Natural Capital and Ecosystem Services Accounting. For the KIP INCA project, we inferred the potential distribution of bumblebees across Europe using species occurrences and their relationships with key environmental drivers.

Environmental drivers were chosen through a combination of ecological criteria (specific for this group) and statistics (see Methods). Species records were made available by the EU-FP7 funded STEP project ('Status and Trends of European Pollinators'), at a spatial accuracy of 10 km. They consisted of validated presence-only bumblebee records, gathered from different data donors in Europe (Suppl. material 7). These records are currently stored in the Atlas Hymenotpera (Rasmont and Francis 2018), a database which also includes records from other taxa of insects and which is constantly updated. We obtained access to bumblebee records for the period between 1970 and 2012, distributed throughout the territories of the European Union (except for Malta and Cyprus, for which the data were not yet available). These territories excluded the Outermost Regions and the Overseas Countries and Territories. The original dataset is described in Rasmont et al. (2015), while an update for the years 2010-2014 is presented in Rasmont and Iserbyt (2014).

We defined:

- occupancy as the number of 10 km grid cells with records for a given bumblebee species;
- record as a database entry for a given bumblebee species, with a spatial accuracy of 10 km.

Acknowledging that the same grid cell can host more than one record of the same species, it follows that the species' occupancy is always equal or less than the number of records for the same species.

Species data

The latitude and longitude of each record (world geodetic system WGS 1984, EPSG:4326) were projected to the LAEA (Lambert azimuthal equal-area projection, EPSG:3035) over a 10 km spatial resolution grid. For the purpose of this work, we used bumblebee records available from 1991 to 2012 inclusive, identified to the level of species. To allow us to obtain robust predictions, we considered all species with an occupancy of at least 25 grid cells (47 species). The list is included in Suppl. material 1, together with the list of species excluded from the analysis.

Environmental predictors

We used four types of environmental predictors:

- Land use / land cover (LULC): percentage cover of 16 continuous variables, derived from the 100 m spatial resolution CORINE Land Cover 2006 accounting layer (version 18.5, Mar. 2017) (Corine LC). This layer was produced by the European Environmental Agency (EEA) through a harmonisation method, allowing the use of the CORINE LC data series to detect changes over time. We chose the CORINE LC hierarchical level 3, the finest one available across the whole area of interest and converted each class to percentage cover within the same 10 x 10 km grid cell of the species data. We then reduced the number of classes to a set of 16, which were ecologically relevant for the taxa. The correspondence with the original CORINE LC classes is listed in Suppl. material 2.
- Climate: we computed monthly averages from daily minimum and maximum temperature and daily total precipitation, using the 25 km spatial resolution E-OBS gridded meteorological data. Since our interest was to have an indication of the climatic pattern, we computed the 12 monthly averages from two decades-data (1991-2012 to match bumblebee records), taking into account leap years. The monthly averages for each of these climatic variables (minimum temperature, maximum temperature and precipitation) served as inputs to compute 19 bioclimatic variables. Bioclimatic variables are biologically relevant variables representing annual trends, seasonality and extreme or limiting environmental factors. For these reasons, they are often used in species distribution modelling. The 19 bioclimatic variables (see WorldClim for their description) can be computed, for instance, with the 'biovars' function of the dismo package in R.
- Topography: we computed mode and standard deviation of elevations within each 10 x 10 km grid cell, using the European Digital Elevation Model (EU-DEM) from Copernicus (25 m spatial resolution, 7 m vertical accuracy). We used these two variables to characterise each 10 km grid cell in terms of its elevation and roughness.
- Others: we computed the 'Average distance from natural and semi-natural areas' (snd_km) within each 10 x 10 km grid cell, rounded to the nearest metreand then converted to kilometre. A number of field-based studies reviewed by Ricketts et al. (2008) and Garibaldi et al. (2011), in fact, prove that increasing distance from semi-

natural areas has a negative impact on a number of pollinator-related outcomes (namely: pollinator richness, visitation rates and stability of pollination services). CORINE LC classes, considered 'natural or semi-natural', are listed in Suppl. material 2.

To minimise multicollinearity between predictors (Guisan and Thuiller 2005), we employed Jolliffe's Principal Component Analysis with the rejection method B4 (Jolliffe 1972, Jolliffe 1973). Details of this procedure and the Pearson's correlation between the predictors are in Suppl. material 3.

In total, 22 predictors were used (Table 1): 16 aggregated LC classes, 4 bioclimatic variables, 1 topographic layer and 1 layer showing the average distance from semi-natural areas.

Table 1.

Environmental predictors used to derive species distribution.

Theme	Definition	Units	Name
Land use / land cover	Agriculture witd natural vegetation	Percent cover	AGNV
	Arable land	Percent cover	AL
	Broad-leaved forest	Percent cover	BF
	Coniferous forest	Percent cover	CF
	Discontinuous urban fabric	Percent cover	DUF
	Green urban areas	Percent cover	GUS
	Heterogeneous agricultural areas	Percent cover	HAG
	Inland waters	Percent cover	IWB
	Inland wetlands	Percent cover	IW
	Mixed forest	Percent cover	MF
	Natural grasslands	Percent cover	NG
	Pastures	Percent cover	PA
	Permanent crops	Percent cover	PC
	Salt marshes	Percent cover	BW
	Scrub vegetation associations	Percent cover	SMH
	Sparsely vegetated areas, including beaches and dunes	Percent cover	BDSV
Climate	Temperature seasonality	Standard deviation *100	TempSeas (bio04)
	Max. temperature of warmest montd	Degree Celsius	MaxTWarmM (bio05)
	Mean temperature of tde wettest quarter	Degree Celsius	MeanTWetQ (bio08)

Theme	Definition	Units	Name
	Precipitation seasonality	Coefficient of variation	RainSeas (bio15)
Topography	Mode of elevations in tde 10-km grid, from tde original DEM	Metre	elmode
Otders	Average distance from natural and semi-natural areas	Kilometre	snd_km

Model settings

Species distribution models were carried out within Maxent (Maximum Entropy Modelling of Species Geographic Distributions, Version 3.4.0, December 2016) (Phillips et al. 2004, Phillips et al. 2006, Phillips et al. 2017). One of the advantages of Maxent is the possibility to use presence-only data, such as the bumblebee records of our study. In addition to the environmental conditions at localities where a species is found, the model requires a random sample from the study region (i.e. from the background). This sample assumes a uniform survey / visiting record over the entire study area. This is seldom the case when species records come from different donors and geographic regions. In case this assumption is violated, the background information should reflect the sample bias. A possible correction is to restrict the selection of the background points to a region with records of similar species (called '**target group**', TG) observed by similar methods (Phillips et al. 2009). We tested for violation of this assumption by comparing the AUC (Area Under the Curve of the Receiver Operating Characteristic) of models based on all records of bumblebees from 1970 to 2012 (and not necessarily identified to the level of species), against the AUC of models based on *n*-time sampling of an equal number of points from the entire study area (referred as '**null models**') (Raes and ter Steege 2007). We found that the average AUCs of models from 100 sets of 500, 1000 and 5000 points drawn from the TG were significantly greater than the average AUCs of an equivalent number of null models from the entire study area (Suppl. material 4). The background localities for the individual SDMs were therefore drawn from within the TG.

We followed the model calibration described in Polce et al. (2013):

- We fitted the data with "Hinge" features, which are base functions for piecewise linear splines (Phillips and Dudík 2008).
- We modified default values of "Prevalence", which is defined as the probability of presence at ordinary occurrence points. When absence data are not available, Maxent assigns it 0.5 (Phillips and Dudík 2008). Prevalence is defined over specific spatial and temporal scales, which should be taken into account particularly when working with pools of species differing in their rarity (Elith et al. 2010). We followed Polce et al. (2013) and empirically computed prevalences, using the number of available records, the number of occurrences and the number of years with non-zero observations over the temporal scale (1991-2012). Each species was then assigned a value for prevalence from 0.1 to 0.5 (Suppl. material 1).
- Additional model settings are listed in Suppl. material 4.

We assessed model predictions using the AUC, which, despite known assumptions and limitations (Austin 2007, Termansen et al. 2006), is commonly used as a threshold-independent measure of model performance within SDMs. With presence-only data such as the pollinators' sightings, the maximum achievable AUC is less than one (Wiley et al. 2003), so standard thresholds for evaluating "goodness of fit" do not apply. As an alternative, we compared the average AUC value of each pollinator distribution model (AUC_{SDM}) with the average AUC value of a set of null models (AUC_{NM}) where species records were replaced by randomly chosen locations (Raes and ter Steege 2007). We expected $AUC_{SDM} > AUC_{NM}$.

At last, model predictions were interpreted as 'Probability of occurrence'.

For each species, two main outputs were considered:

- The average 'Probability of occurrence' across the area of interest, from each of the five model runs.
- The average threshold indicating, for each model run, the species 'Minimum training presence'. This threshold was used to convert the average probability of occurrence (P_{occ}) to 'Presence/Absence' across the study area: if P_{occ}>= Threshold, then presence = 1 or else presence = 0.

These outputs were also used to derive:

- A 'Species richness' map, by summing each species 'Presence/Absence' map.
- An average 'Probability of occurrence' map, from the aggregated set of 47 species: single species P_{occ} maps were summed and their average extracted by dividing it by the 'Species Richness' map. Hence, for each grid cell, the average P_{occ} was based only on the number of species likely to be present.

Results

The assessment of model predictions through the comparison of species-model AUCs and null-model AUCs showed in all cases a performance significantly better than random (Suppl. material 4).

The predicted probability of occurrence for each species, the species richness (i.e. the number of bumblebee species predicted to be present) and the composite probability (average probability of occurrence for those species predicted to be present) are listed in Suppl. material 5 (the "Dataset"), for each pair of geographic coordinates (of the 10 km spatial resolution grid). Low resolution figures derived from these predictions are shown in Suppl. material 6, together with the low resolution figures of the species occupancies used as model inputs. Fig. 1 shows the predicted number of bumblebee species across EU28, Fig. 2 shows the predicted probability of occurrence from species predicted to be present (i.e. for species with $P_{occ} > 0$); in other words, an overall probability of occurrence for bumblebees.

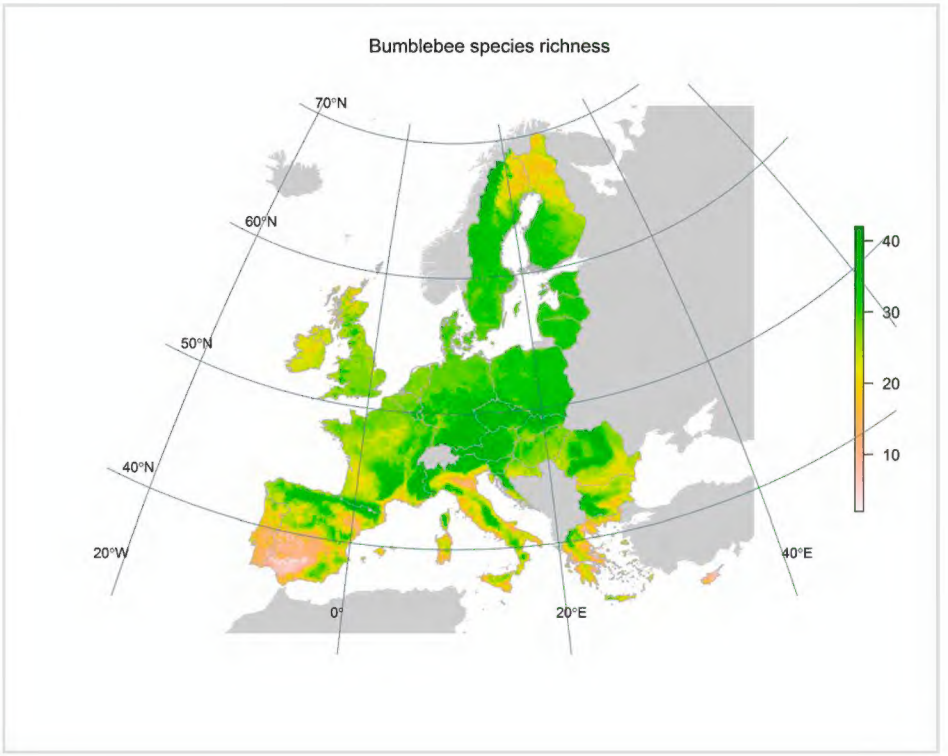

Figure 1.

Predicted number of bumblebee species, derived from the sum of each species presence maps. Presence was defined for each species using an individual threshold applied to the predicted probability of occurrence (see main text for details). Values above the threshold were set to 'presence' (1) and values below it to 'absence' (0). The sum of 'presences' indicate, for each locality on the map, the number of bumblebee species potentially present, i.e. the potential species richness.

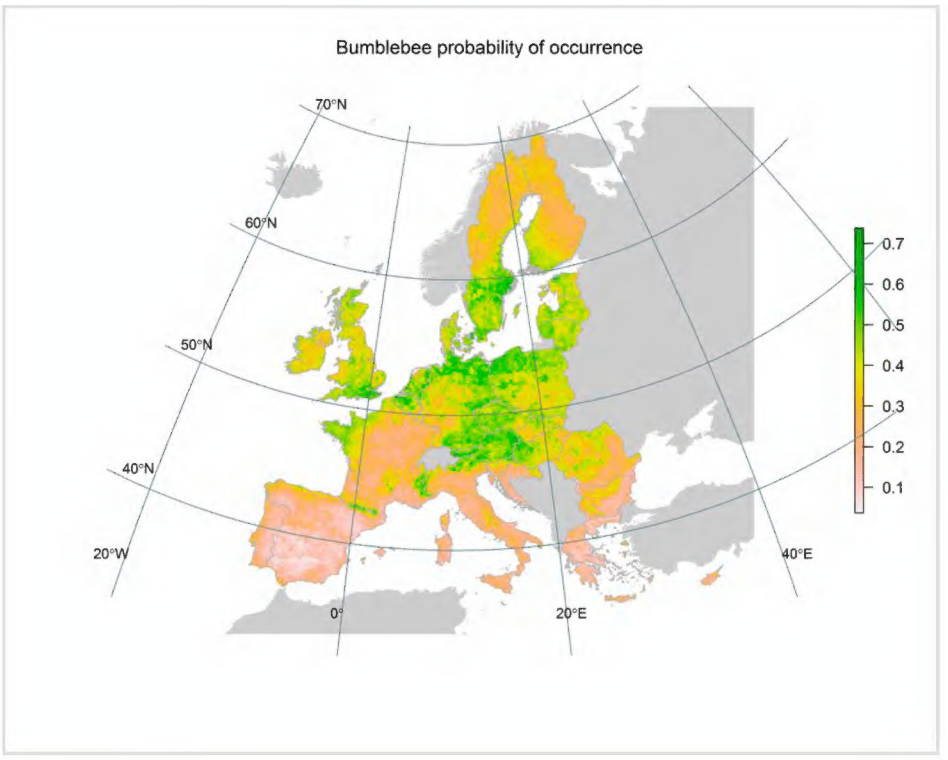

Figure 2.

Predicted probability of occurrence for bumblebees, resulting from the individual species distribution maps. Single species distribution maps were summed and their average extracted by dividing it by the species richness map (Fig. 1). Hence, for each map locality, the average probability of occurrence was based only on the number species predicted to be present.

Dataset

Distribution of bumblebees species across Europe.

A csv table with 43665 rows and 51 columns.

Columns heading: Geographic coordinates (X, Y), predicted probability of occurrence for 47 species of bumblebees (each listed with its species name according to the binomial nomenclature), obtained with Maxent; species richness (number of bumblebee species predicted to be present); composite probability (average probability of occurrence from species predicted to be present).

Geographical coverage: European Union (28 Member States)

Spatial reference system: ETRS89 Lambert Azimuthal Equal Area (epsg projection 3035 - etrs89 / etrs-laea).

Spatial resolution: 10 km

Input data

Species data: bumblebee records from the Atlas Hymenoptera for the period 1991-2012.

Environmental predictors (continuous variables)

- land use and land cover classes (LULC): cover percentage of each class, within the 10 km grid cells
- bio-climatic predictors
- topography
- distance from natural and semi-natural LULC classes

Object name

SM05_BB_Predictions.csv

Format names and versions

With reference to the csv table of Suppl. material 5(the "Dataset"):

X, Y columns: integer numbers of X and Y coordinate in LAEA Spatial Reference System;

Columns from 3 to 49: continuous positive numbers quantifying for each row (i.e. for each locality identified by the X and the Y) the predicted probability of occurrence of individual species of bumblebees (47 in total, each identified by the species name according to the binomial nomenclature);

Column 50: Richness: predicted number of bumblebee species predicted to be present (i.e. with probability of occurrence greater than zero);

Column 51: CompositeProbability: average probability of occurrence, from those species predicted to be present.

Creation dates

May 2018

Dataset creators

Chiara Polce

Dataset contributors

See list of co-authors and Suppl. material 7.

Repository location

This paper (Suppl. material 5)

Publication date

Accepted for publication on 06 September 2018

Re-use potential

Any re-use of these data must acknowledge this source. Any interpretation of results obtained from re-using these data must acknowledge the characteristics of our dataset, which are described within this paper. The authors may be contacted in case of any doubts.

Details for replicability and reproducibility

Input occurrence records used to derive these data are constantly updated; they may be requested from the Atlas Hymenoptera.

Additional information

Occupancy maps are based on records available from 1991 to 2012 inclusive, identified to the level of species and with a spatial accuracy of at least 10 km.

The predicted **probability of occurrence** must be interpreted as the **potential environmental suitability** based on the set of environmental predictors and the input occupancy. Additional records and updated environmental information can lead to different predictions.

The **presence** of a species is derived **from the predicted probability of occurrence, after applying a threshold** to discriminate presence from absence. No common rules exist to choose the threshold. A combination of statistical and ecological criteria might be used. For this study, we adopted '*Minimum training presence*' as a threshold rule, which uses the suitability associated with the least suitable training presence record as the threshold. While this rule might lead to optimistic predictions (also assigning presence to areas at the margins of the species' ecological requirements), it also shows the potential suitability of the environment for the species. This information can be used, for instance, to select target areas for specific pollinator-friendly measures aiming at improving local environmental suitability.

Further research could investigate the possibility of adopting different thresholds for species' presence, based on species' commonality. Such fine-tuning could bring the predicted potential distribution closer to the species' actual distribution.

Disclaimer

The views expressed in this article are personal and do not necessarily reflect an official position of the European Commission.

Acknowledgements

This work could not have been completed without the support of many people and organisations. In particular, we acknowledge Stéphanie Iserbyt, Samantha Bailey, David Baldock, Lucas Baliteau, Renzo Barbattini, Andreas Bertsch, Eduardas Budrys, Frank Burger, Adrien Chorein, Maurizio Cornalba, Graziano Gabriele , Andrej Gogala, Yves Gonseth, Dirk de Graaf, Aljaz Jenic, Dries Laget, Xavier Lair, Anders Nielsen, Frode Ødegaard, Theodora Petanidou, Marino Quaranta, Menno Reemer, Didier Roustide, Peter Sima, Ilkka Teräs, Bernard Vaissière, the ALARM project ("Assessing large scale risks for biodiversity with tested methods", funded by the European Commission 6th Framework Programme, GOCE, -CT-2003-506675) and the STEP project ("Status and trends of European pollinators", funded by the European Commission 7th Framework Programme, 244090).

Author contributions

Pierre RASMONT collected and organised the original species records made available for this work. Data contributors are listed in Suppl. Material 7.

Xavier ROTLLAN-PUIG and Chiara POLCE processed the data and performed model calibration.

Chiara POLCE produced the final models.

Chiara POLCE and Joachim MAES drafted the text and all authors contributed to its final version.

References

- Austin M (2007) Species distribution models and ecological theory: A critical assessment and some possible new approaches. Ecological Modelling 200: 1-19. https://doi.org/10.1016/j.ecolmodel.2006.07.005
- Biesmeijer JC (2006) Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. Science 313 (5785): 351-354. https://doi.org/10.1126/science.1127863
- Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2010) A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17 (1): 43-57. https://doi.org/10.1111/j.1472-4642.2010.00725.x
- Gallai N, Salles J, Settele J, Vaissière B (2009) Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. Ecological Economics 68 (3): 810-821. https://doi.org/10.1016/j.ecolecon.2008.06.014
- Garibaldi LA, Steffan-Dewenter I, Kremen C, Morales JM, Bommarco R, Cunningham SA, Carvalheiro LG, Chacoff NP, Dudenhöffer JH, Greenleaf SS, Holzschuh A, Isaacs R, Krewenka K, Mandelik Y, Mayfield MM, Morandin LA, Potts SG, Ricketts TH, Szentgyörgyi H, Viana BF, Westphal C, Winfree R, Klein AM (2011) Stability of pollination services decreases with isolation from natural areas despite honey bee visits. Ecology Letters 14 (10): 1062-1072. https://doi.org/10.1111/j.1461-0248.2011.01669.x
- Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecology Letters 8 (9): 993-1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x
- IPBES (2016) The assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services on pollinators, pollination and food production. Potts SG, Imperatriz-Fonseca VL, Ngo HT (Eds). Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Bonn, Germany, 552 pp.
- Jolliffe IT (1972) Discarding Variables in a Principal Component Analysis. I: Artificial Data. Applied Statistics 21 (2): 160. https://doi.org/10.2307/2346488
- Jolliffe IT (1973) Discarding Variables in a Principal Component Analysis. II: Real Data. Applied Statistics 22 (1): 21. https://doi.org/10.2307/2346300

- Klein AM, Vaissiere BE, Cane JH, Steffan-Dewenter I, Cunningham SA, Kremen C, Tscharntke T (2007) Importance of pollinators in changing landscapes for world crops. Proceedings of the Royal Society B: Biological Sciences 274 (1608): 303-313. https://doi.org/10.1098/rspb.2006.3721
- Leonhardt SD, Gallai N, Garibaldi LA, Kuhlmann M, Klein AM (2013) Economic gain, stability of pollination and bee diversity decrease from southern to northern Europe. Basic and Applied Ecology 14: 461-471. https://doi.org/10.1016/j.baae.2013.06.003
- Marshall L, Biesmeijer J, Rasmont P, Vereecken N, Dvorak L, Fitzpatrick U, Francis F, Neumayer J, Ødegaard F, Paukkunen JT, Pawlikowski T, Reemer M, Roberts SM, Straka J, Vray S, Dendoncker N (2017) The interplay of climate and land use change affects the distribution of EU bumblebees. Global Change Biology 24 (1): 101-116. https://doi.org/10.1111/gcb.13867
- Nieto A, Roberts SPM, Kemp J, Rasmont P, Kuhlmann M, García Criado M, Biesmeijer JC, Bogusch P, Dathe HH, De la Rúa P, De Meulemeester T, Dehon M, Dewulf A, Ortiz-Sánchez FJ, Lhomme P, Pauly A, Potts SG, Praz C, Quaranta M, Radchenko VG, Scheuchl E, Smit J, Straka J, Terzo M, Tomozii B, Window J, Michez D (2014) European Red List of Bees. Publication Office of the European Union, Luxembourg, 98 pp. [ISBN 978-92-79-44512-5]
- Phillips S, Dudík M, Schapire R (2004) A maximum entropy approach to species distribution modeling. Proceedings of the Twenty-First International Conference on Machine Learning, 655-662 pp.
- Phillips S, Anderson R, Schapire R (2006) Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231-259. https://doi.org/10.1016/j.ecolmodel.2005.03.026
- Phillips S, Anderson R, Dudík M, Schapire R, Blair M (2017) Opening the black box: an open-source release of Maxent. Ecography 40: 887-893. https://doi.org/10.1111/ecog.03049
- Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31: 161-175. https://doi.org/10.1111/j.0906-7590.2008.5203.x
- Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecological Applications 19: 181-197.
- Polce C, Termansen M, Aguirre-Gutiérrez J, Boatman N, Budge G, Crowe A, Garratt M, Pietravalle S, Potts S, Ramirez J, Somerwill K, Biesmeijer J (2013) Species distribution models for crop pollination: a modelling framework applied to Great Britain. PLoS ONE 8 (10): e76308. https://doi.org/10.1371/journal.pone.0076308
- Polce C, Garratt MP, Termansen M, Ramirez-Villegas J, Challinor AJ, Lappage MG, Boatman ND, Crowe A, Endalew AM, Potts SG, Somerwill KE, Biesmeijer JC (2014) Climate-driven spatial mismatches between British orchards and their pollinators: increased risks of pollination deficits. Global Change Biology 20 (9): 2815-2828. https://doi.org/10.1111/gcb.12577
- Potts SG, M Roberts SP, Dean R, Marris G, Brown MA, Jones R, Neumann P, Settele J (2010) Declines of managed honey bees and beekeepers in Europe. Journal of Apicultural Research 49 (1): 15-22. https://doi.org/10.3896/ibra.1.49.1.02

- Raes N, ter Steege H (2007) A null-model for significance testing of presence-only species distribution models. Ecography 30: 727-736. https://doi.org/10.1111/j.2007.0906-7590.05041.x
- Rasmont P, Iserbyt S (2014) Atlas of the European Bees: genus *Bombus*. 3d Edition. STEP Project, Atlas Hymenoptera, Mons, Gembloux. http://www.atlashymenoptera.net/page.asp?id=169
- Rasmont P, Franzen M, Lecocq T, Harpke A, Roberts S, Biesmeijer K, Castro L, Cederberg B, Dvorak L, Fitzpatrick U, Gonseth Y, Haubruge E, Mahe G, Manino A, Michez D, Neumayer J, Odegaard F, Paukkunen J, Pawlikowski T, Potts S, Reemer M, Settele J, Straka J, Schweiger O (2015) Climatic risk and distribution atlas of European bumblebees. BioRisk 10: 1-236. https://doi.org/10.3897/biorisk.10.4749
- Rasmont P, Francis F (2018) Atlas Hymenoptera. Université de Mons, Université de Liège. http://www.atlashymenoptera.net/. Accessed on: 2018-6-01.
- Ricketts TH, Regetz J, Steffan-Dewenter I, Cunningham SA, Kremen C, Bogdanski A, Gemmill-Herren B, Greenleaf SS, Klein AM, Mayfield M, Morandin LA, Ochieng' A, Viana BF (2008) Landscape effects on crop pollination services: are there general patterns? Ecology Letters 11 (5): 499-515. https://doi.org/10.1111/j.1461-0248.2008.01157.x
- Termansen M, McClean C, Preston C (2006) The use of genetic algorithms and Bayesian classification to model species distributions. Ecological Modelling 192: 410-424. https://doi.org/10.1016/j.ecolmodel.2005.07.009
- Vallecillo S, La Notte A, Polce C, Zulian G, Alexandris N, Ferrini S, Maes J (2018) Ecosystem services accounting: Part I - Outdoor recreation and crop pollination, EUR 29024 EN. Publications Office of the European Union, Luxembourg. [ISBN 978-92-79-77333-4] https://doi.org/10.2760/619793
- Wiley EO, McNyset K, Peterson T, Robins R, Stewart A (2003) Niche modeling perspective on geographic range predictions in the marine environment using a machine-learning algorithm. Oceanography 16 (3): 120-127. https://doi.org/10.5670/oceanog.2003.42
- Williams IH (1994) The dependence of crop production within the European Union on pollination by honey bees. Agricultural Zoology Reviews 6: 229-257.
- Zulian G, Maes J, Paracchini M (2013) Linking land cover data and crop yields for mapping and assessment of pollination services in Europe. Land 2 (3): 472-492. https://doi.org/10.3390/land2030472

Supplementary materials

Suppl. material 1: SM01_Species.doc doi

Authors: Chiara Polce
Data type: Text and table
Brief description: List of bumblebee species for which a distribution model was derived and list of those excluded. Occupancy refers to the presence of at least one species' record for each 10-km grid cell. Prevalence is a parameter used by Maxent and was derived following Polce et al. (2013).
Filename: SM01_Species.doc - Download file (89.00 kb)

Suppl. material 2: SM02_LandUseLandCover.doc doi

Authors: Chiara Polce
Data type: Text and tables
Brief description: Table S II.1: Land use / land cover (LULC) predictors with reference to the original CORINE land cover classes.
Table S II.2: CORINE LC classes defining semi-natural areas.
Filename: SM02_LandUseLandCover.doc - Download file (73.50 kb)

Suppl. material 3: SM03_CorrelationBetweenPredictors.doc doi

Authors: Chiara Polce
Data type: Text and tables
Brief description: Table S III.2 Pearson's correlation between all bioclimatic and topographic predictors.
Table S III.3: Pearson's correlation between bioclimatic and topographic predictors selected to minimise multicollinearity.
Filename: SM03_CorrelationBetweenPredictors.doc - Download file (86.00 kb)

Suppl. material 4: SM04_ModelSettingsAndPerformance.doc doi

Authors: Chiara Polce
Data type: Text and Charts
Brief description: AUC and model settings
Filename: SM04_ModelSettingsAndPerformance.doc - Download file (126.00 kb)

Suppl. material 5: SM05_BB_Predictions.csv doi

Authors: Chiara Polce
Data type: Table
Brief description: csv table with values for:
- x and y geographic coordinates in Lambert Azimuthal Equal Area, with a spatial resolution of 10 km;
- species-level probability of occurrence for 47 bumblebee species;
- species richness (number of species predicted to be present);
- composite probability (average probability of occurrence for those species predicted to be present).
Filename: SM05_BB_Predictions.csv - Download file (9.49 MB)

Suppl. material 6: SM06_Bumblebee_distribution.pdf doi

Authors: Chiara Polce
Data type: PDF file showing distribution maps
Brief description: Pages 2 to 48: two maps on each page, showing input species occupancies and predicted probability of occurrence, for 47 species of bumblebees.
Occupancy maps are based on records available from 1991 to 2012 inclusive, identified to the level of species and with a spatial accuracy of at least 10 km.
The predicted probability of occurrence must be interpreted as the potential environmental suitability based on the set of environmental predictors and the input occupancy. Additional records and updated environmental information can lead to different predictions.

Page 49: two maps showing the overall probability of occurrence for those species predicted to be present in each grid cell and predicted species richness derived therefrom.
The presence of a species is derived from the predicted probability of occurrence, after applying a threshold to discriminate presence from absence. No common rules exist to choose the threshold. A combination of statistical and ecological criteria might be used. For this study, we adopted 'Minimum training presence' as a threshold rule, which uses the suitability associated with the least suitable training presence record as the threshold. While this rule might lead to optimistic predictions (also assigning presence to areas at the margin of the species' ecological requirements), it also shows the potential suitability of the environment for the species. This information can be used, for instance, to select target areas for specific pollinator-friendly measures aiming at improving local environmental suitability.
See main text for additional details and for the data sources.
Filename: SM06_Bumblebee_distribution.pdf - Download file (6.01 MB)

Suppl. material 7: SM07_DataContributors_20120922.xls doi

Authors: Pierre Rasmont
Data type: MO Excel table
Brief description: A table-like spreadsheet listing primary sources of bumblebee records used to derive the predictions presented in this paper.
Filename: SM07_DataContributors_20120922.xls - Download file (52.50 kb)